

1 40735/JEC/D465

SYSTEM AND METHOD FOR CREATING, STORING, AND UTILIZING  
COMPOSITE IMAGES OF A GEOGRAPHIC LOCATION

5

CROSS-REFERENCE TO RELATED APPLICATION(S)

This application claims the benefit of U.S. provisional  
application entitled SYTEM FOR CREATING AND USING A VISUAL RECORD  
OF A GEOGRAPHIC LOCATION COUPLED WITH POSITION INFORMATION,  
10 Provisional Application No. 60/238,490, filed on October 6, 2000  
the content of which is incorporated herein by reference.

FIELD OF THE INVENTION

15 This invention relates to visual databases, specifically  
to the creation and utilization of visual databases of  
geographic locations.

BACKGROUND OF THE INVENTION

20 There exist methods in the prior art for creating visual  
databases of geographic locations. However, such databases are  
of limited use due to the method of acquiring the imagery as  
well as the kind of imagery acquired. One particular method  
involves the taking of individual photographs of the location  
25 and electronically pasting the photographs on a polygonal mesh  
that provide the framework for a three-dimensional (3D)  
rendering of the location. This method, however, is time  
consuming and inefficient for creating large, comprehensive  
databases covering a substantial geographic area such as an  
30 entire city, state, or country.

Another method uses video technology to acquire the  
images. The use of video technology, especially digital video  
technology, allows the acquisition of the image data at a higher  
35 rate, reducing the cost involved in creating the image  
databases. For example, the prior art teaches the use of a  
295554-1

1 40735/JEC/D465

vehicle equipped with a video camera and a Global Positioning System (GPS) to collect image and position data by driving  
5 through the location. The video images are later correlated to the GPS data for indexing the imagery. Nevertheless, such a system is still limited in its usefulness due to the lack of additional information on the imagery being depicted.

The nature of the acquired imagery also limits the  
10 usefulness of such a system. A single image acquired by the video camera contains a narrow field of view of a locale (e.g. a picture of a single store-front) due to the limited viewing angle of the video camera. This narrow field of view provides  
15 little context for the object/scene being viewed. Thus, a user of such an image database may find it difficult to orient himself or herself in the image, get familiar with the locale, and navigate through the database itself.

One way to increase the field of view is to use a shorter  
20 focal length for the video camera, but this introduces distortions in the acquired image. Another method is to increase the distance between the camera and the buildings being filmed. However, this may not be possible due to the limit on the width of the road and constructions on the opposite side of  
25 the street.

The prior art further teaches the dense sampling of images of an object/scene to provide different views of the object/scene. The sampling is generally done in two dimensions  
30 either within a plane, or on the surface of an imaginary sphere surrounding the object/scene. Such a sampling, however, is computationally intensive and hence cumbersome and inefficient in terms of time and cost.

Accordingly, there is a need for a system and method for  
35 creating a visual database of a comprehensive geographic area in a more time and cost efficient manner. Such a system should

not require the reconstruction of 3D scene geometry nor the  
dense sampling of the locale in multiple dimensions.  
5 Furthermore, the images in the database should provide a wider  
field of view of a locale to provide context to the objects  
being depicted. The database should further correlate the  
images with additional information related to the geographic  
location and objects in the location to further enhance the  
10 viewing experience.

#### SUMMARY OF THE INVENTION

The present invention addresses and alleviates the above-  
mentioned deficiencies associated with the prior art. More  
15 particularly, the present invention is directed to a computer-  
implemented system and method for synthesizing images of a  
geographic location to generate composite images of the  
location. The geographic location may be a particular street  
20 in a geographic area with the composite images providing a view  
of the objects on each side of the street.

According to one aspect of the invention, an image  
recording device moves along a path recording images of objects  
along the path. A GPS receiver and/or inertial navigation  
25 system provides position information of the image recording  
device as the images are being acquired. The image and position  
information is provided to a computer to associate each image  
with the position information.

The computer synthesizes image data from the acquired  
30 images to create a composite image depicting a view of the  
objects from a particular location outside of the path.  
Preferably, the composite image provides a field of view of the  
location that is wider than the field of view provided by any  
35 single image acquired by the image recording device.

In another aspect of the invention, the path of the camera is partitioned into discrete segments. Each segment is preferably associated with multiple composite images where each composite image depicts a portion of the segment. The composite images and association information are then stored in an image database.

In yet another aspect of the invention, the image database contains substantially all of the static objects in the geographic area allowing a user to visually navigate the area from a user terminal. The system and method according to this aspect of the invention identifies a current location in the geographic area, retrieves an image corresponding to the current location, monitors a change of the current location in the geographic area, and retrieves an image corresponding to the changed location. A map of the location may also be displayed to the user along with information about the objects depicted in the image.

#### DESCRIPTION OF THE DRAWINGS

FIG. 1 is a schematic block diagram of a data acquisition and processing system for acquiring image and position data used to create composite images of a geographic location;

FIG. 2 is an illustration of a composite image created by the data acquisition and processing system of FIG. 1;

FIG. 3 is a high-level flow diagram of the steps exercised by the data acquisition and processing system of FIG. 1 in creating and storing the composite images;

FIG. 4 is a flow diagram for synchronizing image sequences with position sequences of a recording camera according to one embodiment of the invention;

FIG. 5 is a flow diagram of an alternative embodiment for synchronizing image sequences with position sequences of a recording camera;

FIG. 6 is a block diagram of the data acquisition and processing system of FIG. 1 allowing a real-time synchronization of image and position data;

FIG. 7 is another embodiment for synchronizing image sequences with position sequences of a recording camera;

FIG. 8 is a flow diagram for segmenting and labeling a camera trajectory;

FIG. 9 is an illustration of a trajectory in a single camera scenario;

FIG. 10 is a flow diagram for creating a composite image of a segment of a camera's path;

FIG. 11 is a flow diagram for identifying and retrieving an optical column from an acquired image according to one embodiment of the invention;

FIG. 12 is a flow diagram for identifying and retrieving an optical column from an acquired image according to an alternative embodiment of the invention;

FIG. 13 is an illustration of an exemplary street segments table including street segments in a camera's trajectory;

FIG. 14 is an illustration of an exemplary image coordinates table for associating composite images with the street segments in the street segments table of FIG. 13;

FIG. 15 is an illustration of an exemplary segment block table for allowing an efficient determination of a segment that is closest to a particular geographic coordinate;

FIG. 16 is an illustration of an exemplary graphical user interface for allowing the user to place requests and receive information about particular geographic locations;

FIG. 17 is a flow diagram of a process for obtaining image and location information of an express street address; and

5 FIG. 18 is a flow diagram of the process for obtaining image and location information of a location selected from a map.

#### DETAILED DESCRIPTION OF THE INVENTION

10 FIG. 1 is a schematic block diagram of a data acquisition and processing system for acquiring and processing image and position data used to create composite images of a geographic location. The composite images are created by synthesizing individual image frames acquired by a video camera moving through the location and filming the objects in its view. The composite  
15 images may depict an urban scene including the streets and structures of an entire city, state, or country. The composite images may also depict other locales such as a zoo, national park, or the inside of a museum, allowing a user to visually  
20 navigate the locale.

The data acquisition and processing system includes one or more image recording devices preferably taking the form of digital video cameras 10 moving along a trajectory/path and recording images on the trajectory on digital videotapes 12.  
25 Other types of acquisition devices may also be used in combination to, or in lieu of, the digital cameras 10, such as analog cameras. Furthermore, the video images may be recorded on optical, magnetic, or silicon video tapes, or on any other known types of storage devices that allow random access of  
30 particular image frames and particular video pixels within the image frames.

The data acquisition and processing system further includes a GPS receiver 16 for receiving position information from a set  
35 of GPS satellites 18 as the cameras 10 move along the trajectory. An inertial navigation system 20 including one or more

295554-1 6

accelerometers and/or gyroscopes also provides position information to the data acquisition and processing system. When  
5 the inertial navigation system 20 is used in conjunction with the GPS receiver 16, a more accurate calculation of the position information may be produced.

In an alternative embodiment, position information is acquired using devices other than the inertial navigation system  
10 20 and/or the GPS receiver 16, such as via computer-vision-based algorithms that compute positions using video information from the video cameras 10.

The video cameras 10 provide to a multiplexer 22 a frame number and time information for each image acquired via a  
15 communication link 24 preferably taking the form of a LANC<sup>TM</sup> port. The GPS receiver 16 and inertial navigation system 20 also provide position information to the multiplexer 22 via communication links 26a, 26b, preferably taking the form of RS-  
20 232 ports. The multiplexer 22 in turn transmits the received frame number, time information, and position data to a data acquisition computer 34 via a communication link 30, which also preferably takes the form of an RS-232 port. The computer 34 stores in a trajectory database 36 the position data from the GPS  
25 receiver 16 and/or inertial navigation system 20 along with the frame number and time information from the video cameras 10. This information is then used by a post-processing system 38 to create the composite images.

The post-processing system 38 preferably includes a post-  
30 processing computer 28 in communication with a video player 39. The computer 28 preferably includes a video acquisition card for acquiring and storing the image sequences as the video player 39 plays the videotapes 12 of the acquired images. The computer 28  
35 includes a processor (not shown) programmed with instructions to take the image and position data and create one or more composite  
295554-1

images for storing into an image database 32. The image database 32 is preferably a relational database that resides in a mass storage device taking the form of a hard disk drive or drive array. The mass storage device may be part of the computer 28 or a separate database server in communication with the computer.

In an alternative embodiment, the images are transferred directly to the data acquisition computer 34 as the images are being recorded. In this scenario, the computer 34 is preferably equipped with the video acquisition card and includes sufficient storage space for storing the acquired images. In this embodiment, the data acquisition computer 34 preferably contains program instructions to create the composite images from the acquired images.

In general terms, a composite image of a particular geographic location is created by using at least one video camera 10 recording a series of video images of the location while moving along a path. In the one camera scenario, the camera 10 is moved twice on the same path but in opposite directions to film the objects on both sides of the path. Movement to the camera 10 is provided by a base, platform, or motor vehicle moving at an average speed of preferably about 20 miles/hour to ensure a sufficient resolution in the resulting images. Video cameras with higher sampler rates may allow for faster acquisition speeds.

Preferably, the data acquisition and processing system uses four cameras 10 mounted on top of a moving motor vehicle. Two side cameras face each side of the path for filming objects viewed from the each side of the vehicle. A front and back cameras allow the filming of the objects viewed from the front and back of the vehicle. The front and back cameras may be equipped with fish-eye lens for providing a wide-angle view of the path. A person skilled in the art should recognize,

295554-1



however, that additional cameras may be used to film the objects from different viewing directions. For example, a duodecahedron  
5 of cameras may be used to record the objects from all viewing directions. Furthermore, the side cameras need not face directly to the side of the street, but may face a slightly forward or backward direction to provide a look up or down of the path.

10 As the images acquired by the cameras 10 are recorded on the videotapes 12, the frame number and time associated with the images are transferred to the data acquisition computer 34. The images recorded on the videotapes 12 are later transferred to the  
15 post-processing computer 28 for further processing. Once the images are received, the computer 28 stores the image data in its memory in its original form or as a compressed file using one of various well-known compression schemes, such as MPEG.

As the camera 10 moves along the path and records the  
20 objects in its view, the GPS receiver 16 computes latitude and longitude coordinates using the information received from the set of GPS satellites 18 at selected time intervals (e.g. one sample every two seconds). The latitude and longitude coordinates indicate the position of the camera 10 during the recording of  
25 a particular image frame. The GPS satellite 18 also transmits to the GPS receiver 16 a Universal Time Coordinate (UTC) time of when the coordinates were acquired. The GPS receiver 16 is preferably located on the vehicle transporting the camera 10 or on the camera itself. The GPS data with the position sequences  
30 and UTC time information is then transferred to the computer 34 for storing in the trajectory database 36.

In addition to the position information provided by the GPS receiver 16, the inertial navigation system 20 also provides  
35 acceleration information to the computer 34 for independently deriving the position sequence of the camera 10. Specifically,  
295554-1

the one or more accelerators and gyroscopes in the inertial navigation system 20 monitor the linear and rotational acceleration rates of the camera 10 and transfer the acceleration data to the computer 34. The computer 34 integrates the acceleration data to obtain the position of the camera 10 as a function of time. The computer 34 preferably combines the position derived from the acceleration information with the GPS position data to produce a more accurate evaluation of the position of the camera 10 at particular instances in time.

The post-processing computer 28 uses the image and position sequences to synthesize the acquired images and create composite images of the location that was filmed. The composite images preferably provide a wider field of view of the location than any single image frame acquired by the camera 10. In essence, the composite images help provide a panoramic view of the location.

FIG. 2 is an illustration of a composite image 40 created from the image frames 42 acquired by the camera 10 while moving along an x-axis 58 direction. In creating the composite image 40, the computer assumes a fictitious camera 44 located behind the actual camera 10 and identifies optical rays 46 originating from the fictitious camera. The location of the fictitious camera 44 depends on the desired field of view of the location being filmed. The further away the fictitious camera is placed from the objects along the x-axis 58, the wider its field of view of the objects.

The computer also identifies optical rays 48 originating from the actual camera 10. For each optical ray 46 from the fictitious camera 44, the computer 28 selects an acquired image frame 42 that includes a corresponding optical ray 48 originating from the actual camera 10. Image data from each selected image frame 42 is then extracted and combined to form the composite image. Preferably, the image data extracted from each image

frame is an optical column that consists of a vertical set of pixels. The composite image is preferably created on a column-by-column basis by extracting the corresponding optical column from each image frame. Thus, to create a column  $P_i$  50 in the composite image 40, the computer locates an image frame 42a that was acquired when the camera 10 was located at  $X_i$  52. To locate this image frame 42a, the computer uses the GPS data and/or data from the inertial navigation system 20 to identify a time  $T_i$  54 at which the camera 10 was in the location  $X_i$  52. Once the image frame 42a is identified, the computer 28 extracts the optical column 56 with an index  $(P_i/N)*M$ , where  $N$  is the total number of columns in the composite image 40 and  $M$  is the number of columns in the image frame 42a. The composite image 40 is stored in the image database 32, preferably in JPEG format, and associated with an identifier identifying the particular geographic location depicted in the image. Furthermore, close-ups and fish-eye views of the objects are also extracted from the video sequences using well-known methods, and stored in the image database 32. The unused data from the acquired images is then preferably deleted from the computer's memory.

FIG. 3 is a high-level flow diagram of the steps exercised by the data acquisition and processing system in creating and storing the composite images. In step 60, the camera 10 acquires a series of images of a particular geographic location. At the same time, the GPS receiver 16 and/or inertial navigation system 20 acquires the position of the camera 10 while the images are being acquired. Because the time associated with the position information (position time) is likely to differ from the times of acquisition of the video images (video time), the computer 28, in step 62, synchronizes the image sequence with the position sequence. The synchronization is preferably a post-processing

step that occurs after the image and position sequences have been acquired.

5 In step 64, the computer 28 segments the trajectory taken by the recording camera 10 into multiple segments and labels each segment with identifying information about the segment. For example, if the camera traverses through various streets, the computer 28 segments the trajectory into multiple straight street  
10 segments and associates each street segment with a street name and number range. In step 66, the computer 28 generates a series of composite images depicting a portion of each segment, and in step 68, stores each composite image in the image database 32 along with the identifying information of the segment with which  
15 it is associated.

FIG. 4 is a more detailed flow diagram of step 62 for synchronizing the image sequences with the position sequences of the recording camera according to one embodiment of the  
20 invention. Although the process illustrated in FIG. 4 assumes that the position data is GPS data, a person skilled in the art should recognize that a similar process may be employed to synchronize the images to positions calculated using the inertial navigation system 20.

25 The process starts, and in step 70, a user of the system selects a landmark in the image sequence that appears in at least two distinct video frames. This indicates that the landmark was recorded once while the camera 10 was moving on one direction on the path, and again while the camera was moving in an opposite  
30 direction on the same path. The landmark may be, for example, a tree in a lane divider.

In step 72, a time interval  $T$  is measured in the image sequence between the two passings of the landmark. In step 74,  
35 the computer 28 uses the GPS data to compute a function for determining the time interval between successive passes of any  
295554-1 12

point along the path. The function is then used to find, for each point  $x$  on the path, a time of return  $Tr(x)$  which measures  
 5 the lapse of time between the two passings of each point. In step 76, a point is identified for which  $Tr(x)=T$ . The identified point provides the GPS position of the landmark and hence, a GPS time associated with the landmark. Given the GPS time, a difference between the GPS time and the video time associated  
 10 with the landmark may be calculated for synchronizing any image frame acquired at a particular video time to the GPS position of the camera at a particular GPS time.

FIG. 5 is a flow diagram of an alternative embodiment for synchronizing the image sequences with GPS position information. As in FIG. 4, the process illustrated in FIG. 5 also identifies,  
 15 in step 80, a landmark in the image sequence that appears in at least two distinct image frames. In step 82, a time phase is initialized to an arbitrary value using the camera time stamp. In step 84, the computer 28 measures the distance traveled  
 20 between the two points on the path that correspond to the time instants in the image sequence where the landmark is seen from the two sides of the path. In step 86, an inquiry is made as to whether the distance has been minimized. If the answer is NO,  
 25 the time phase is modified in step 88, and the process returns to step 84 where the distance is measured again.

In another embodiment, the synchronization does not occur as a post-production process, but occurs in real-time as the image and position sequences are acquired. FIG. 6 is a block  
 30 diagram of a portion of the data acquisition and processing system of FIG. 1 allowing a real-time synchronization of image and position data. A UTC clock generator 90 provides to the data acquisition computer 34 the UTC time associated with each GPS  
 35 position of the recording camera 10 as the camera moves along the path. The video time produced by a camera clock 92 is also

provided to the data acquisition computer 34 via the communications port 24. A UTC card 94 on the computer 34 correlates the video time to the UTC time. Thus, the video image acquired at the video time may be correlated to the GPS position of the camera during the recording of the image.

FIG. 7 is yet another embodiment for synchronizing the image sequences with the GPS position of the recording camera 10. In step 100, the post-processing computer 28 computes the temporal variation in the image values (i.e. optical flow) of the bottom pixel rows in the image sequence. Thus, the average velocity of each of the pixels in the row may be obtained. In step 102, the tangential velocity of the camera 10 is obtained from the GPS reading. The average velocity for the computed pixels is directly proportional to the vehicle tangential velocity. Thus, in step 104, the time phase between the position and video sequences may be determined as a time delay maximizing the alignment of local maxima and local minima between the average pixel velocity and the vehicle tangential velocity. This time phase is then read out in step 106.

FIG. 8 is a more detailed flow diagram of step 64 of FIG. 3 for segmenting the trajectory followed by one or more recording cameras 10 and labeling the segments with an identifier. In the one camera scenario, the camera is moved along the path making a right turn at each intersection until a block 112 has been filmed, as is illustrated in FIG. 9. The camera then moves to a second block 114 to film the objects on that block. Thus, a particular path 110 adjoining the two blocks 112, 114 is traversed twice on opposite directions allowing the filming of the objects on each side of the path.

In step 120, the post-processing computer 28 segments the camera's trajectory into straight segments by detecting the points of maximum curvature (i.e. where the turns occur). In

1 40735/JEC/D465

this regard, the latitude and longitude coordinates provided by the GPS receiver 16 are converted into two-dimensional Mercator coordinates according to well-known methods. A spline interpolation is then obtained from the two-dimensional Mercator coordinates and the resulting spline function is parameterized in arc-length. The computer 28 obtains a new sampling of the coordinates from the spline function by uniformly sampling the coordinates in an arc-length increment of about one meter while detecting the points in the new sequence where a turn was made. The place where a turn occurs is assumed to be the place of an intersection of two segments.

15 Preferably, the computer 28 performs a singular value decomposition computation according to well-known methods to detect the turns. In this regard, the computer selects an observation window containing N sample points that is moved along the spline for calculating an index indicative of the overall direction (i.e. alignment) of the points in the window. The higher the index, the less aligned the points, and the more likely that the camera was making a turn at those points. The points are least aligned at the center of a turn, and thus, the computer selects as a turn coordinate a point in the observation window where the index is at a local maximum. The computer 28 gathers all the points whose indexes correspond to local maxima and stores them into an array of turn coordinates.

30 In step 122, the computer 28 determines the center of an intersection by grouping the turn coordinates into clusters where turns that belong to the same cluster are turns made on the same intersection. An average of the turn coordinates belonging to the same cluster is then calculated and assigned as the intersection coordinate.

35 The endpoints of each straight segment are identified based on the calculated intersection coordinates. In this regard, an  
295554-1 15

1 40735/JEC/D465

intersection coordinate at the start of the segment is identified and assigned to the segment as a segment start point (the "From" intersection coordinate). An intersection coordinate at the end of the segment is also identified and assigned to the segment as a segment end point (the "To" intersection coordinate).

In the scenario where at least two side cameras are utilized, each camera films the objects on each side of the path during the first pass on the path. Thus, unlike the single camera scenario where a turn is made at each intersection to move the camera along the same path twice but in opposite directions, a turn is not made at each intersection in the two camera scenario. Therefore, instead of identifying the points of maximum curvature for determining the intersection coordinates, the intersection coordinates are simply detected by tracking the GPS data and identifying where the segments orthogonally intersect.

The computer 28 associates the calculated segments with information obtained from a geographic information database 128 and stores it into a segments table as is described in further detail below. In the scenario where composite images of a city are created, the geographic information database 128 includes a map of the city where the endpoints of each street segment on the map are identified by latitude and longitude information. The database 128 further includes a street name and number range for each street segment on the map. Such databases are commercially available from third parties such as Navigation Technologies and Etak, Inc.

In correlating the segments of the camera's trajectory with the segments in the geographic information database 128, the computer, in step 124, determines the correspondences between the "From" and "To" coordinates calculated for the trajectory segment with intersection coordinates of the segments in the database.



1 40735/JEC/D465

The computer 28 selects the segment in the geographic information database 128 which endpoints are closest to the computed "From" and "To" coordinates, as the corresponding segment.

In step 126, the computer labels each trajectory segment with information that is associated with the corresponding segment in the database 128. Thus, if each segment in the database 128 includes a street name and number, this information is also associated with the trajectory segment.

FIG. 10 is a more detailed flow diagram of step 66 of FIG. 3 for creating a composite image of a segment of the camera's path according to one embodiment of the invention. In step 130, the computer 28 computes the arc length coordinate  $X_c$  of the center of the segment which is then set as the center of the composite image. In step 132, the computer identifies the optical rays 46 (FIG. 2) originating from the fictitious camera 44 by computing an array of equidistant positions  $X_1, X_2, \dots, X_n$  along the camera's trajectory, centered around  $X_c$ . The number of computed positions preferably depend on the number of optical columns that are to be created in the composite image.

In step 134, the computer 28 uses the position information obtained from the GPS receiver 16 and/or inertial navigation system 20 to map each position  $X_i$  on the trajectory to a position time  $T_i$ . Thus, if GPS data is used to determine the camera's position, each position  $X_i$  of the camera 10 is mapped to a UTC time.

In step 136, the computer 28 uses the time phase information computed in the synchronization step 62 of FIG. 3 to convert the position times to video times. For each identified video time, the computer 28, in step 138, identifies an associated image frame and extracts a column of RGB pixel values from the frame corresponding to the optical rays 46 originating from the fictitious camera 44. In step 140, the column of RGB

295554-1

pixel values are stacked side by side to generate a single image bitmap forming the composite image.

5 FIG. 11 is a more detailed flow diagram of step 138 for identifying and retrieving a column of RGB pixel values for a particular video time  $T_i$  according to one embodiment of the invention. In step 150, the computer 28 identifies an image frame with frame index  $F_i$  acquired at time  $T_i$ . Because the image  
10 frames are acquired at a particular frame rate (e.g. one frame every  $1/30$  seconds), there may be a particular time  $T_i$  for which an image frame was not acquired. In this scenario, the frame closest to time  $T_i$  is identified according to one embodiment of the invention.

15 In step 152, the current position of the image sequence is set to the image frame with index  $F_i$ , and the frame is placed into a frame buffer. In step 154, a column in the image frame with an index  $i$  is read out from the frame buffer.

20 FIG. 12 is a flow diagram of an alternative embodiment for identifying and retrieving a column of RGB pixel values for a particular video time  $T_i$ . If an image frame was not acquired at exactly time  $T_i$ , the computer, in step 160, identifies  $2*N$  image frames that are closest to time  $T_i$ , where  $N \geq 1$ . Thus, at least  
25 two image frames closest to time  $T_i$  are identified. In step 162, the computer computes an optical flow field for each of the  $2*N$  image frames and in step 164, infers the column of RGB values for a column  $i$  at time  $T_i$ . In the situation where only two image frames are used to compute the optical flow, the computer  
30 identifies for each pixel in the first image frame a position  $X_1$  and a position time  $T_1$ . A corresponding pixel in the second frame is also identified along with a position  $X_2$  and a position time  $T_2$ . The position  $X'$  of each pixel at time  $T_i$  is then  
35 computed where  $X' = X_1 + ((T_i - T_1)/(T_2 - T_1)) * (X_2 - X_1)$ . Given the position of each pixel at time  $T_i$ , a new frame that

corresponds to time  $T_i$  may be inferred. The computer 28 may then extract the column of RGB values from the new frame for a column  
5 i.

Preferably, the computer 28 creates multiple composite images at uniform increments (e.g. every 8 meters) along a segment. In the scenario where the composite images are created for street segments, the composite images depict the view of the  
10 objects on each side of the street. The composite images are then stored in the image database 28 along with various tables that help organize and associate the composite images with street segment information.

According to one embodiment of the invention, the image  
15 database 32 includes composite images of a geographic area which together provide a visual representation of at least the static objects in the entire area. Thus, if the geographic area is a particular city, the composite images depict the city on a  
20 street-by-street basis, providing a visual image of the buildings, stores, apartments, parks, and other objects on the streets. The system further includes an object information database with information about the objects being depicted in the composite images. If the geographic area being depicted is a  
25 city, the object information database contains information about the structures and businesses on each city street. In this scenario, each record in the object information database is preferably indexed by a city address.

FIG. 13 is an illustration of an exemplary street segments  
30 table 170 including the street segments in the camera's trajectory. The table 170 suitably includes multiple entries where each entry is a record specific to a particular street segment. A particular street segment record includes an index  
35 identifying the street segment (segment ID) 172 as well as the corresponding street name 174 obtained from the geographic  
295554-1

1 40735/JEC/D465

information database 128 (FIG. 12). A particular street segment  
record also includes the direction of the street (North, South,  
5 East, or West) 176 with respect to a main city street referred  
to as a city hub. The direction information generally appears  
in an address in front of the street name. A city, state, and/or  
country fields may also be added to the table 170 depending on  
the extent of the geographic area covered in the image database  
10 32.

A street segment record includes the endpoint coordinates  
178 of the corresponding street segment in the geographic  
information database 128. An array of segment IDs corresponding  
to street segments adjacent to the segment start point are  
15 identified and stored in field 180 along with the direction in  
which they lie with respect to the start point (e.g. North,  
South, East, or West). Similarly, an array of segment IDs  
corresponding to street segments adjacent to the segment end  
20 point are also identified and stored in field 182. These  
segments are also ordered along the direction in which they lie.

In addition to the above, a street segment record includes  
a distance of the start of the trajectory segment from the city  
hub 184. The city hub generally marks the origin of the streets  
from which street numbers and street directions (North, South,  
25 East, or West) are determined. Street numbers are generally  
increased by two at uniform distances (e.g. every 12.5 feet)  
starting from the hub. Thus the distance from the hub allows a  
computation of the street numbers on the street segment. In a  
30 one camera situation where each segment is traversed twice, the  
distance from the hub is computed for each camera trajectory. In  
a multiple camera scenario, however, only one distance is  
computed since the camera traverses the segment only once.

35 Also included in a street segment record is a length of the  
trajectory segment. Such a length is computed for each  
295554-1 20

trajectory in a one camera 10 scenario because the movement of the camera 10 is not identical during the two traversals of the segment.

Each street segment record 170 further includes an offset value 188 for each side of the street. The offset is used to correct the street numberings computed based on the distance information. Such a computation may not be accurate if, for instance, there is an unusually wide structure on the segment that is erroneously assigned multiple street numbers because it overlaps into the area of the next number assignment. In this case, the offset is used to decrease the street numbers on the segment by the offset value.

FIG. 14 is an illustration of an exemplary image coordinates table 200 for associating the composite images with the street segments in the street segments table 170. The image coordinates table 200 includes a plurality of composite image records where each record includes a segment ID 202 of the street segment being depicted in the composite image. In addition, each composite image record includes information of the side of the street segment 204 being depicted. For example, the side may be described as even or odd based on the street numbers on the side of the street being depicted. Each composite image entry also includes a distance from the segment origin to the center Xc of the composite image 206 indicating the position along the street segment for which the image was computed. The distance information is used to retrieve an appropriate composite image for each position on the street segment.

FIG. 15 is an illustration of an exemplary segment block table 210 for allowing an efficient determination of a segment ID that is closest to a particular geographic coordinate. In this regard, the geographic area depicted in the image database 32 is preferably partitioned into a grid of square blocks where

1 40735/JEC/D465

each block includes a certain number of street segments. The blocks are assigned block labels preferably corresponding to the center longitude and latitude coordinates of the block. The block labels are stored in a block label field 212. Associated with each block label are segment IDs 214 corresponding to the street segments in the block. Given the coordinates of a particular geographic location, the block closest to the given coordinates may be identified to limit the search of street segments to only street segments within the block.

In a particular use of the system, a user places inquiries about a location in a geographic area depicted in the image database 32. For example, the user may enter an address of the location, enter the geographic coordinates of the location, select the location on a map of the geographic area, or specify a displacement from a current location. Preferably, the user has access to a remote terminal that communicates with a host computer to service the user requests. The host computer includes a processor programmed with instructions to access the image database 32 in response to a user request and retrieve composite images about the particular location. The processor is also programmed with instructions to access the geographic and object information databases to retrieve maps and information on the businesses in the geographic area. The retrieved data is then transmitted to the requesting remote user terminal for display thereon.

The remote user terminals may include personal computers, set-top boxes, portable communication devices such as personal digital assistants, and the like. The visual component of each remote user terminal preferably includes a VGA or SVGA liquid-crystal-display (LC) screen, an LED display screen, or any other suitable display apparatus. Pressure sensitive (touch screen) technology may be incorporated into the display screen so that

the user may interact with the remote user terminal by merely touching certain portions of the screen. Alternatively, a keyboard, keypad, joystick, mouse, and/or remote control unit is provided to define the user terminal's input apparatus.

Each remote user terminal includes a network interface for communicating with the host computer via wired or wireless media. Preferably, the communication between the remote user terminals and the host computer occurs over a wide area network such as the Internet.

In an alternative embodiment of the invention, the image, geographic information, and object information databases reside locally at the user terminals in a CD, DVD, hard disk drive, or any other type of mass storage media. In this embodiment, the user terminals include a processor programmed with instructions to receive queries from the user about a particular geographic location and retrieve composite images and associated information in response to the user queries.

FIG. 16 is an illustration of an exemplary graphical user interface (GUI) for allowing the user to place requests and receive information about particular geographic locations. The GUI includes address input fields 220 allowing the user to enter the street number, street name, city and state of the particular location he or she desires to view. Actuation of a "See It" button 222 causes the user terminal to transmit the address to the host computer to search the image and geographic location databases 32, 128 for the composite image and map corresponding to the address. Furthermore, the host computer searches the object information database to retrieve information about the objects depicted in the composite image.

The retrieved composite image and map are respectively displayed on the display screen of the requesting user terminal in a map area 226 and an image area 224. The map is preferably

1 40735/JEC/D465

centered around the requested address and includes a current  
location cursor 228 placed on a position corresponding to the  
5 address. The current location identifier 228 may, for instance,  
take the shape of an automobile.

The composite image displayed on the image area 224  
provides a view of a side of the street (even or odd) based on  
the entered street number. The user may obtain information about  
10 the objects being visualized in the composite image by actuating  
one of the information icons 234 above the image of a particular  
object. In displaying the information icons 234, a range of  
street addresses for the currently displayed image is computed.  
The listings in the object information database with street  
15 numbers that fall inside the computed range are then selected and  
associated with the information icons 234 displayed on top of the  
image of the object.

If the objects are business establishments, the information  
20 displayed upon actuating the information icons 234 may include  
the name, address, and phone number 236 of the establishment.  
This information is preferably displayed each time the user  
terminal's cursor or pointing device is passed above the icon.  
In addition, if the establishment is associated with a particular  
25 Web page, the information icon 234 functions as a hyperlink for  
retrieving and displaying the Web page, preferably on a separate  
browser window.

The user may obtain a close-up view of a particular object  
30 in the composite image by selecting the object in the image. A  
close-up view of the object is then obtained by computing the  
distance of the selected object from the origin of the street  
segment where the object lies. The location on the segment of  
the closest close-up image is computed and retrieved from the  
35 image database 32. The close-up image is then provided in the  
image area 224 or in a separate browser window.



A "Switch View" button 230 allows the user to update the current composite image providing a view of one side of the street with a composite image of the other side of the street. In a multiple camera scenario, each actuation of the "Switch View" button 230 provides a different view of the street. The current view is preferably identified by a direction identifier (not shown) on or close to the current location identifier 228. For instance, one side of the current location identifier 228 may be marked with a dot or an "X" to identify the side of the street being viewed. Alternatively, an arrow may be placed near the current location identifier 228 to identify the current viewing direction.

The composite image is also updated as the user navigates through the streets using the navigation buttons 232. From a current location, the user may choose to navigate north, south, west, and east by actuating the corresponding buttons. Preferably, only the navigation buttons corresponding to the possible direction of motions from the current position are enabled. As the user actuates one of the buttons, the current position is incremented by a predetermined amount, currently set to eight meters, to the next available composite image on the current or adjacent segment. The image area 224 is then updated with the new composite image.

FIG. 17 is a flow diagram of the process executed by the host computer for obtaining image and location information of an express street address entered in the address input fields 220. A similar process is executed by the user terminal in the embodiment where the location and image information are stored locally at the user terminal.

The process starts, and in step 240, the user requests information about a particular street address by entering the address in the address input fields 220. In step 242, the

1 40735/JEC/D465

address is transmitted to the host computer preferably over a wide area network such as the Internet. In step 244, a query is  
5 run on the host computer to locate the street segment index in the street segment table 170 (FIG. 13) corresponding to the requested address. In this regard, the computer searches the street segment table 170 for street segments that match the desired street name 174 as well as a city, state, or country if  
10 applicable. For each street segment matching the street name, the computer computes the starting street number on that segment preferably based on the following formula:

15 
$$\text{Start Number} = (\text{round}((\text{Distance from Hub} + \text{Offset})/K)*2)$$

The distance from the hub 184 and offset 188 values are obtained from the street segment table 170. The value K is the distance assumed between any two street numbers on the segment.

20 The ending street number on the street segment is also calculated using a similar formula:

$$\text{End Number} = (\text{round}((\text{Distance from Hub} + \text{Offset} + \text{length})/K)*2)$$

25 The length is the length 186 value obtained from the street segment table 170.

Once the start and end street numbers are calculated for a particular street segment, the computer determines whether the requested street number lies within the start and end street  
30 numbers. If it does, the computer returns the corresponding segment ID 172. Furthermore, the computer determines the distance of the requested street number from the start of the street segment for determining the position of the street number  
35 on the street segment.

In step 246, the host computer transmits the query result to the requesting user terminal along with a map of the input location retrieved from the geographic information database 128. In step 248, the requesting user terminal downloads from the host computer a record from the street segments table 170 corresponding to the identified street segment. Furthermore, the user terminal also retrieves the computed start and end street numbers of the street segment, a list of computed composite images for both sides of the street segment in the image coordinates table 200 (FIG. 14), and information of the objects visible on the street segment in the object information database.

In step 250, the user terminal downloads a composite image for the appropriate side of the street from the host computer that has a distance from the origin of the street segment to the center of the composite image 206 (FIG. 14) that is closest to the distance of the desired street number from the origin. The display on the user terminal is then updated in step 252 with the new location and image information.

FIG. 18 is a flow diagram of the process executed by the host computer for obtaining image and location information of a particular location selected on the map displayed in the map area 226. A similar process is executed by the user terminal in the embodiment where the location and image information are stored locally at the user terminal.

The process starts, and in step 260, the user requests information about a particular street address by selecting a location on the map. In step 262, the map coordinates are converted from screen coordinates to geographic location coordinates (x,y) and transmitted to the host computer preferably over the Internet. In step 244, a query is run on the host computer to locate the street segment index in the street segment table 170 (FIG. 13) corresponding to the

geographic location coordinates. In this regard, the computer searches the segment block table 210 (FIG. 15) for a block that includes the street segment corresponding to the input location. In order to locate such a block, the computer rounds the identified geographic location coordinates based preferably on the size of the block. The rounded (x,y) coordinates may thus be represented by  $((\text{round}(x/B)) * B, (\text{round}(y/B) * B))$ , where B is the length of one of the block sides. The computer then compares the rounded number to the coordinates stored in the block label field 212 and selects the block with the label field 212 equal to the rounded coordinate. Once the appropriate block is identified, the computer proceeds to retrieve the segment IDs 214 associated with the block. The geographic coordinates of the desired location are then compared with the endpoint coordinates of each street segment in the block for selecting the closest street segment.

In step 266, the segment ID of the closest street segment is returned to the user terminal. Additionally, a new map of the desired location may be transmitted if the previous map was not centered on the desired location.

In step 268, the requesting user terminal downloads from the host computer a street segment record in the street segments table 170 corresponding to the identified street segment. The user terminal also retrieves the calculated start and end street numbers of the street segment, a list of computed composite images for both sides of the street the segment in the image coordinates table 200 (FIG. 14), and information of the objects visible on the street segment in the object information database.

In step 270, the user terminal downloads the composite image corresponding to the geographic coordinates of the input location. To achieve this, the geographic coordinates are

converted to a distance along the identified street segment. The user terminal downloads a composite image that has a distance from the origin of the street segment to the center of the composite image 206 (FIG. 14) that is closest to the distance of the input location from the origin. The display on the user terminal is then updated in step 272 with the new location and image information.

10 Although this invention has been described in certain specific embodiments, those skilled in the art will have no difficulty devising variations which in no way depart from the scope and spirit of the present invention. For example, the composite images may be made into streaming video by computing the composite images at small increments along the path (e.g. every 30 cm). Furthermore, the composite images may be computed at several resolutions by moving the fictitious camera 44 (FIG. 2) closer or further away from the path to decrease or increase its field of view and provide the user with different zoom levels of the image.

Variation may also be made to correct any distortions in the perspective of the composite image along the vertical y-axis direction. The extraction of the optical columns from the acquired image frames may introduce such a distortion since the sampling technique used along the horizontal x-axis direction is not applied along the y-axis. Such a distortion may be corrected by estimating the depth of each pixel in the composite image using optical flow. The aspect ratio of each pixel may be adjusted based on the distance of the object visualized in the pixel. The distortion may also be corrected by acquiring images from an array of two or more video cameras 10 arranged along the vertical y-axis in addition to the cameras in the horizontal axis.

The described method of generating composite images also  
 relies on an assumption that the camera's trajectory is along  
 5 a straight line. If this is not the case and the vehicle  
 carrying the camera makes a lane change, makes a turn, or passes  
 over a bump, the choice of the optical column extracted from a  
 particular image frame may be incorrect. The distortion due to  
 such deviations from a straight trajectory may, however, be  
 10 corrected to some degree using optical flow to detect such  
 situations and compensate for their effect.

It is therefore to be understood that this invention may be  
 practiced otherwise than is specifically described. Thus, the  
 present embodiments of the invention should be considered in all  
 15 respects as illustrative and not restrictive, the scope of the  
 invention to be indicated by the appended claims and their  
 equivalents rather than the foregoing description.

20

25

30

35